

**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU  
PREHRAMBENO-TEHNOLOŠKI FAKULTET OSIJEK**

**Iva Vasilj**

**KEMOMETRIJSKI MODEL ZAVISNOSTI SVOJSTAVA KVALITETE  
PŠENIČNOG ZRNA I BRAŠNA**

**DIPLOMSKI RAD**

**Osijek, listopad, 2016**

## TEMELJNA DOKUMENTACIJSKA KARTICA

DIPLOMSKI RAD

Sveučilište Josipa Jurja Strossmayera u Osijeku  
Prehrambeno-tehnološki fakultet Osijek  
Zavod za procesno inženjerstvo  
Katedra za modeliranje, optimiranje i automatizaciju  
Franje Kuhača 20, 31000 Osijek, Hrvatska

### Diplomski sveučilišni studij Procesno inženjerstvo

**Znanstveno područje:** Biotehničke znanosti

**Znanstveno polje:** Prehrambena tehnologija

**Nastavni predmet:** Modeliranje operacija i procesa

**Tema rada** je prihvaćena na IX. redovitoj sjednici Fakultetskog vijeća Prehrambeno-tehnološkog fakulteta Osijek u akademskoj godini 2015./2016. održanoj 28. lipnja 2016.

**Mentor:** prof. dr. sc. *Damir Magdić*

**Komentor:** dr. sc. *Daniela Horvat*, znan. savj.

### Kemometrijski model zavisnosti svojstava kvalitete pšeničnog zrna i brašna

*Iva Vasilj*, 289-DI

#### Sažetak:

Kemometrijskim metodama su određene povezanosti između svojstava pšeničnog zrna i brašna od 24 kultivara. Kultivari su uzgajani na poljima Poljoprivrednog instituta Osijek tijekom deset godina (2005.-2014.). Analizirana svojstva su: prinos pšenice (Y), hektolitarska masa (HL), masa tisuću zrna pšenice (TKW), udio proteina (P), udio vlažnog glutena (WG), gluten indeks (GI), sedimentacijska vrijednost (SED), broj padanja (FN) i izbrašnjavanje (FY). Na izmjerenom setu podataka provedena je sljedeća deskriptivna statistička analiza: određivanje srednjih vrijednosti, medijana, minimalne i maksimalne vrijednosti, standardne devijacije, koeficijenta varijabilnosti i njihovih međusobnih korelacijskih koeficijenata. Provedene su i sljedeće kemometrijske analize: analiza glavnih komponenti (PCA), klaster analiza (CA) i regresijska analiza metodom najmanjih kvadrata (PLSR). Analiza glavnih komponenti korištena je za smanjenje broja varijabli, klaster analiza korištena je za pokazivanje veza među svojstvima, a regresijska analiza metodom najmanjih kvadrata korištena je za izradu prediktivnih modela. Rezultati pokazuju da je na temelju snažnih korelacija moguće smanjiti broj varijabli i sa manje od devet svojstava opisati varijabilnost seta podataka. Predloženim prediktivnim matematičkim modelima moguće je korištenjem izmjerenih vrijednosti izračunavati vrijednosti preostalih svojstava s točnošću većom od 90%.

**Ključne riječi:** pšenično zrno, brašno, svojstva, povezanost, kemometrijske metode

**Rad sadrži:** 42 stranica  
11 slika  
15 tablica  
16 literaturnih referenci

**Jezik izvornika:** hrvatski

#### Sastav Povjerenstva za ocjenu i obranu diplomskog rada i diplomskog ispita:

- |                                                |               |
|------------------------------------------------|---------------|
| 1. Izv.prof. dr. sc. <i>Sandra Budžaki</i>     | predsjednik   |
| 2. prof. dr. sc. <i>Damir Magdić</i>           | član-mentor   |
| 3. dr. sc. <i>Daniela Horvat</i> , znan. savj. | član-komentor |
| 4. izv. prof. dr. sc. <i>Ana Bucić-Kojić</i>   | zamjena člana |

**Datum obrane:** 10. listopada 2016.

**Rad je u tiskanom i elektroničkom (pdf format) obliku pohranjen u** Knjižnici Prehrambeno-tehnološkog fakulteta Osijek, Franje Kuhača 20, Osijek.

## BASIC DOCUMENTATION CARD

GRADUATE THESIS

University Josip Juraj Strossmayer in Osijek  
Faculty of Food Technology Osijek  
Department of Process Engineering  
Sub department of Modelling, Optimisation and Automation  
Franje Kuhača 20, HR-31000 Osijek, Croatia

### Graduate program Process Engineering

**Scientific area:** Biotechnical sciences

**Scientific field:** Food technology

**Course title:** Modeling of operation and process

**Thesis subject** was approved by the Faculty of Food Technology Osijek Council at its session no. IX. held on July 28, 2016.

**Mentor:** *prof. Damir Magdić*, PhD

**Co-mentor:** *Daniela Horvat*, PhD, sci. adv.

### Chemometric dependence model of quality of wheat grain and flour

*Iva Vasilj*, 289-DI

### Summary:

Chemometric methods were used to determine connection among properties of wheat grain and flour produced from 24 cultivars. Cultivars were breed on Agricultural institute Osijek in period of ten years (2005-2014). Following properties were analysed: yield (Y), hectolitre mass (HL), thousand kernel weight (TKW), protein content (P), wet gluten content (WG), gluten index value (GI), sedimentation value (SED), falling number (FN) and flour yield (FY). Following descriptive statistical analysis was applied on experimental set of data: mean value, median, minimum and maximum, standard deviation, coefficient of variability and coefficient of correlation. Following chemometric analysis were applied: principal component analysis (PCA), cluster analysis (CA) and partial least square method (PLSR). Principal component analysis was used for reduction of number of variables, cluster analysis was used for enlightening connections among properties while partial least square regression analysis was used for designing predictive mathematical models. Results show that strong correlations can be used for decreasing number of variables and describing variability of data set with less than nine properties. Using proposed predictive mathematical models based on experimental data makes possible calculating values of rest of properties with precision bigger than 90%.

**Key words:** wheat grain, flour, properties, connection, chemometric methods

**Thesis contains:** 42 pages  
11 figures  
15 tables  
16 references

**Original in:** Croatian

### Defense committee:

1. *assoc. prof. Sandra Budžaki*, PhD
2. *prof. Damir Magdić*, PhD
3. *Daniela Horvat*, PhD, sci. adv.
4. *assoc. prof. Ana Bucić-Kojić*, PhD

Chair person  
Supervisor  
Member  
Stand-in

**Defense date:** October 10, 2016

**Printed and electronic (pdf format) version of thesis is deposited in** Library of the Faculty of Food Technology  
Osijek, Franje Kuhača 20, Osijek.

*Na krilima svog anđela nema ničega što ne možeš biti, učiniti ili imati...*

*Posvećeno bratu Kristijanu-Titiju*

*Zahvaljujem se prof. dr. sc. Damiru Magdiću na predloženoj temi diplomskog rada te na stručnoj pomoći, savjetima i razumijevanju tijekom izrade rada.*

*Zahvaljujem se dr. sc. Danieli Horvat i dipl. ing. Kristini Lutrov s Poljoprivrednog instituta Osijek na pomoći pri izradi diplomskog rada.*

*Veliko hvala obitelji i prijateljima na potpori i razumijevanju tijekom mog studiranja.*

## Popis oznaka, kratica i simbola

PCA	eng. <i>Principal components analysis</i> Analiza glavnih komponenti
CA	eng. <i>Cluster analysis</i> Klaster analiza
PLS	eng. <i>Partial least squares regression</i> Regresija metodom najmanjih kvadrata
Y	eng. <i>Yield</i> prinos ( $\text{kg ha}^{-1}$ )
HL	eng. <i>Hectolitre</i> hektolitarska masa (kg)
TKW	eng. <i>thousand kernel weight</i> masa tisuću zrna (g)
P	eng. <i>protein</i> udio bjelančevina (%)
SED	eng. <i>sedimentation</i> sedimentacijska vrijednost ( $\text{cm}^3$ )
WG	eng. <i>wet gluten</i> vlažni gluten (%)
GI	eng. <i>gluten index</i> gluten indeks
FN	eng. <i>falling number</i> broj padanja (s)
FY	eng. <i>flour yield</i> izbrašnjavanje (%)

## Sadržaj

<b>1. UVOD</b>	<b>1</b>
<b>2. TEORIJSKI DIO</b>	<b>3</b>
<b>2.1. PŠENICA</b>	<b>4</b>
2.1.1. Pšenično zrno	4
2.1.2. Fizikalne osobine kakvoće pšenice	5
2.1.3. Kemijske osobine kakvoće pšenice	5
2.1.4. Prehrambena vrijednost pšenice	5
<b>2.2. PŠENIČNO BRAŠNO</b>	<b>5</b>
2.2.1. Kemijski sastav brašna	6
<b>2.3. TEHNOLOŠKA KAKVOĆA BRAŠNA</b>	<b>6</b>
2.3.1. Karakteristike pšeničnog brašna u pekarstvu	6
2.3.2. Udio i kakvoća glutena	7
<b>2.4. DESKRIPTIVNA STATISTIKA</b>	<b>8</b>
<b>2.5. KEMOMETRIJSKE METODE</b>	<b>10</b>
2.5.1. Analiza glavnih komponenti (PCA)	10
2.5.2. Klaster analiza	11
2.5.3. Regresijska analiza metodom najmanjih kvadrata (PLSR)	11
<b>3. EKSPERIMENTALNI DIO</b>	<b>12</b>
<b>3.1. ZADATAK</b>	<b>13</b>
<b>3.2. MATERIJALI I METODE</b>	<b>13</b>
3.2.1. Određivanje hektolitarske mase	13
3.2.2. Određivanje apsolutne mase	13
3.2.3. Određivanje udjela proteina	13
3.2.4. Određivanje sedimentacijske vrijednosti po Zeleny-u	14
3.2.5. Određivanje vlažnog glutena	14
3.2.6. Određivanje amilolitičke aktivnosti brašna metodom broja padanja po Hagberg-Pertenu	15
3.2.7. Statistička obrada podataka	15
<b>4. REZULTATI</b>	<b>17</b>
<b>4.1. DESKRIPTIVNA STATISTIKA</b>	<b>18</b>
<b>4.2. KEMOMETRIJSKE METODE</b>	<b>23</b>
4.2.1. Analiza glavnih komponenti (PCA)	23
4.2.2. Klaster analiza (CA)	26
4.2.3. Regresijska analiza metodom najmanjih kvadrata (PLSR)	28
<b>5. RASPRAVA</b>	<b>33</b>
<b>6. ZAKLJUČCI</b>	<b>38</b>
<b>7. LITERATURA</b>	<b>41</b>

## **1. UVOD**

Pšenica je jednogodišnja biljka vrste *Triticum* koja se ubraja u porodicu *Poaceae*.

Zdrave, zrele biljke si obično zlatnožute boje, a razlikuju se prema visini (60-120 cm), što zavisi od količine vlage, dužine dnevnog svjetla i količine gnojiva tijekom uzgoja i od genetskih osobina svake vrste (Ugarčić-Hardi, 1999).

Brašno se dobiva mljevenjem očišćenih zrna pšenice iz endosperma pšeničnog zrna, te o njegovom sastavu ovisi kakvoća brašna. Po kemijskom sastavu brašno sadrži: proteine, škrob, celulozu, pentozane, lipide, mineralne tvari, vodu i topljive šećere.

U ovom radu ispitano je 9 svojstava kvalitete pšeničnog zrna i brašna od 24 kultivara pšenice tijekom 10 godina (2005.-2014.).

Na izmjerenom setu podataka provedene su deskriptivna statistika, te kemometrijske metode. Cilj ovog rada bio je smanjiti tj. reducirati broj izvornih varijabli, klasificirati, te predvidjeti različite modele korištenjem kemometrijskih metoda.



## **2. TEORIJSKI DIO**

## 2.1. PŠENICA

Prema pronađenim zapisima i nalazima utvrđeno je kako je pšenica poznata više od 10000 godina, kada je uzgajana u Iraku, Maloj Aziji, Kini i Egiptu. Prije 5000 godina uzgajana je u istočnom dijelu Europe, a nakon otkrića Amerike i Australije počeo je uzgoj pšenice na tim kontinentima (Pomeranz, 2000; Ačkar, 2010).

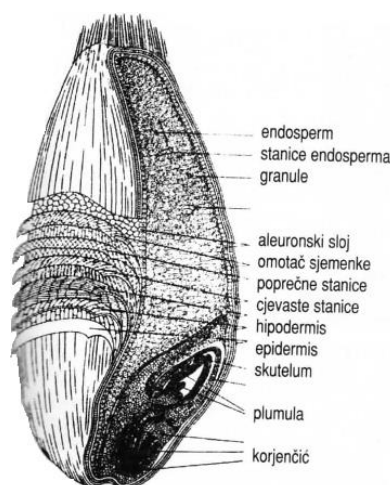
Pšenica je druga žitarica po uzgoju u svijetu, iza kukuruza. Ona je jednogodišnja biljka vrste *Triticum* koja se ubraja u porodicu *Poaceae*, a može biti ozima i jara.

Zdrave, zrele biljke si obično zlatnožute boje, a razlikuju se prema visini (60-120 cm), što zavisi od količine vlage, dužine dnevnog svjetla i količine gnojiva tijekom uzgoja i od genetskih osobina svake vrste (Ugarčić-Hardi, 1999).

Prinos pšenice uvjetovan je različitim agroekološkim uvjetima kao i agrotehničkim mjerama poput vlažnosti i temperature zraka, plodnosti tla, kvalitetnom gnojidbom, itd.

### 2.1.1. Pšenično zrno

Zrno je plod pšenice, a može biti različite krupnoće (krupno, srednje i sitno). U klasu se razvije oko 30-40 zrna (**Slika 1**). Zrno se sastoji od omotača, klice (najmanji, ali biološki najvažniji dio jer se u njoj nalaze svi budući organi biljke) te endosperma (čini najveći dio oko 86% ukupne mase zrna, a u njemu su smještene pričuve hranjivih tvari).



**Slika 1** Uzdužni presjek zrna pšenice (Kljusurić, 2000.)

### **2.1.2. Fizikalne osobine kakvoće pšenice**

Važnije fizikalne osobine zrna pomoću kojih se ocjenjuje kvaliteta pšenice su hektolitarska masa, masa tisuću zrna, specifična masa, veličina i oblik zrna, te staklavost i boja zrna.

Hektolitarska masa pšenice je pokazatelj prinosa brašna, predstavlja masu u jednom hektolitru izraženu u kg. Na hektolitarsku masu utječu oblik, veličina, gustoća zrna, sadržaj vlage u zrnu i primjese.

Apsolutna masa pšenice iznosi tisuću prosječnih zrna pšenice, zavisi od veličine zrna (Ugarčić-Hardi, 1999).

### **2.1.3. Kemijske osobine kakvoće pšenice**

Kemijske osobine zrna su maseni udio vode, maseni udio i kakvoća proteina, maseni udio masti, maseni udio pepela i celuloze.

Udio vode u zrnu kreće se u granicama 10 – 14%. Sadržaj vode iznad 15% je nepovoljan zbog otežanog skladištenja zrna. Maseni udio masti je 1,5 – 2%, najvećim dijelom u klici. Pri složenoj meljavi klica se odvaja pa se brašno može duže čuvati. Celuloza se nalazi u pšeničnoj ljusci s udjelom oko 2 – 3% mase zrna (Kljusurić, 2000).

Sadržaj proteina, najvažnijeg sastojka zrna pšenice ovisi o brojnim čimbenicima, ponajprije vrsti i kultivaru, klimatskim uvjetima, tipu tla i njegovim kemijskim svojstvima (plodnost i agrotehnika).

### **2.1.4. Prehrambena vrijednost pšenice**

Kemijski sastav zrna pšenice ovisi od kultivara, uvjeta uzgoja i kakvoće tla. Promjene se očitavaju u količini i kakvoći proteina, ugljikohidrata, masti, mineralnih tvari i vitamina prisutnih u zrnu pšenice.

## **2.2. PŠENIČNO BRAŠNO**

Mljevenjem očišćene pšenice dobiva se pšenično brašno, primjenom odgovarajućih strojeva uz što veće iskorištenje i minimalni utrošak energije. Primjenom jednostavnog i ekonomičnog procesa prerade i dobre kakvoće pšenice dobiva se brašno bolje kakvoće.

Mljevenjem pšeničnog zrna ljuska se odvaja od endosperma. Mljevenjem ljuske dobije se posije, mljevenjem endosperma brašno (Kljusurić, 2000).

### **2.2.1. Kemijski sastav brašna**

Pšenično brašno dobije se mljevenjem pšenice; dobije se brašno raznih tipova (različito po sastavu i tehnološkim osobinama) ovisno o postotku izbrašnjavanja. Postotak izbrašnjavanja se definira kao odnos udjela brašna određene kakvoće i upotrijebljene pšenice. Pri povećanju postotka izbrašnjavanja pšenice, u brašnu se povećava udio mineralnih tvari, proteina, lipida, enzima, vitamina i sirovih vlakana, a opada udio brašna.

Brašno se sastoji od vode, mineralnih tvari, škroba, proteina, topljivih šećera, celuloze i lipida. Brašno sadrži i enzime i vitamine. Najvažniji enzimi u brašnu su dijestaze, proteaze, lipaze, oksidaze i dr. Od vitamina najzastupljeniji su provitamin A, nikotinska kiselina, tiamin (B1), riboflavin (B2), i tokoferol (E) (Đaković 1980).

## **2.3. TEHNOLOŠKA KAKVOĆA PŠENIČNOG BRAŠNA**

Kakvoća brašna ovisi o vrsti i kultivaru pšenice, te načinu meljave. Ovisno o namjeni i načinu meljave brašna mogu biti raznih tipova. Brašno relativno dobre kakvoće nije pogodno za sve namjene, stoga se za tehnološku kakvoću brašna u pekarstvu upotrebljava izraz pecivost.

Postupak prerade pšenice odvija se u dvije faze: priprema i mljevenje. Priprema zrna pšenice uključuje uklanjanje primjesa i kondicioniranje. Zatim slijedi mljevenje zrna, koje se vrši postepeno u više koraka. U prvom koraku se zrno razdvaja na sastavne dijelove, nakon toga slijedi usitnjavanje krupice i odvajanje endosperma od dijelova ljuske. U zadnjem koraku krupica se melje do završnog proizvoda, brašna.

### **2.3.1. Karakteristike pšeničnog brašna u pekarstvu**

Karakteristike koje definiraju pecivost pšeničnog brašna su:

- Sposobnost razvoja plina je količina plina (CO<sub>2</sub>) koje brašno može razviti u određenom vremenskom periodu pri fermentaciji tijesta. Ovisi o slijedećim čimbenicima: količini

slobodnog šećera i dekstrina u brašnu, koncentraciji enzima, temperaturi, pH, vlažnosti tijesta, količini kvasca, količini kisika.

- Sposobnost zadržavanja plina ovisi o udjelu i kakvoći proteina u brašnu. Tijesto s velikom sposobnošću zadržavanja plina će se dobiti ako proteini (glijadin i glutenin) imaju sposobnost stvaranja ljepka dobre kakvoće.
- Sposobnost vezivanja vode ovisi o količini prisutnih proteina tj. glutena, o tipu i jačini brašna, načinu vođenja procesa. Jakost brašna je sposobnost brašna da u procesu miješanja upije određenu količinu vode, razlikuju se snažna i slaba brašna. Snažna brašna pri miješanju tijesta upijaju veliku količinu vode, te tijesto zadržava svoje fizikalne osobine tijekom zamjesa i fermentacije. Takav proizvod je porozan, velikog volumena i loptast. Slabo brašno upije malu količinu vode, te tijesto ne zadržava svoje fizikalne osobine, postaje meko, ljepljivo i razmazuje se. Gotov proizvod je malog volumena.

### **2.3.2. Udio i kakvoća glutena**

U pšeničnom zrnu proteini su glavna komponenta za određivanje kakvoće brašna (Kljusurić, 2000). Proteini su polipeptidni spojevi sastavljeni od aminokiselina. Gluten je rezervni protein endosperma zrna pšenice koja čini 80-85% ukupnih proteina brašna. Glijadini i glutenini su glavne komponente glutena.

Veliki utjecaj na tehnološku kakvoću brašna ima udio i kakvoća glutena. Tijekom zamjesa, gluten nastaje kao netopljiva bjelančevina brašna, dodatkom vode glijadin i glutenin bubre, te sljepljivanjem tvore mrežnu strukturu. Bitna svojstva mrežne strukture su njezina plastičnost i elastičnost. Prilikom zamjesa bitno je vrijeme zamjesa, te temperatura. Produženim mehaničkim djelovanjem razorila bi se struktura glutena.

Metode za određivanje kakvoće glutena su određivanje sedimentacijske vrijednosti i određivanje gluten indeksa.

## 2.4. DESKRIPTIVNA STATISTIKA

Deskriptivna ili opisna statistika temelji se na potpunom obuhvatu statističkog skupa, čiju masu podataka organizirano prikuplja, odabire, grupira, prezentira i interpretira dobivene rezultate analize.

Artemitička sredina (eng. arithmetic mean) niza podataka  $x_1, x_2, \dots, x_n$  iz varijable X definirana je izrazom:

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

gdje je:  $\bar{x}_n$  artimetička sredina,  $n$  ukupni broj podataka,  $\sum_{i=1}^n x_i$  suma svih produkata.

Medijan je vrijednost statističkog obilježja koja niz dijeli na dva jednaka dijela.

Varijanca i standardna devijacija spadaju u grupu mjera raspršenosti podataka. One karakteriziraju raspršenost podataka oko artimetičke sredine. Varijanca niza izmjerenih vrijednosti  $x_1, x_2, \dots, x_n$  varijable X definirana je izrazom:

$$\bar{s}_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \quad (2)$$

gdje je:  $\bar{s}_n^2$  varijanca,  $\sum_{i=1}^n (x_i - \bar{x}_n)^2$  suma svih kvadrata razlike podataka i srednje vrijednosti.

Standardna devijacija je kvadratni korijen varijance, tj.

$$\bar{s}_n = \sqrt{\bar{s}_n^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2} \quad (3)$$

gdje je:  $\bar{s}_n$  standardna devijacija.

Koeficijent varijabilnosti izračunava se kao udio standardne devijacije u srednjoj vrijednosti:

$$\frac{\bar{s}_n}{\bar{X}_n} \times 100[\%] \quad (4)$$

i objašnjava koliko je odstupanje minimalne i maksimalne vrijednosti u skupu podataka od srednje vrijednosti.

Koeficijent korelacije jedna je numerička karakteristika dvodimenzionalnog slučajnog vektora koja može poslužiti za analizu zavisnosti među njegovim komponentama.

Neka je  $(X, Y)$  dvodimenzionalan slučajni vektor kojemu svaka komponenta ima varijancu. Koeficijent korelacije je broj definiran izrazom:

$$\rho_{XY} = \frac{E(X-\mu)(Y-v)}{\sigma_X \sigma_Y} \quad (5)$$

gdje su

$$\mu = EX, v = EY, \sigma_X = \sqrt{VarX}, \sigma_Y = \sqrt{VarY}$$

O koeficijentu korelacije valja znati sljedeće činjenice:

- $\rho_{XY} \in [-1, 1]$
- ako su  $X$  i  $Y$  nezavisne slučajne varijable tada je  $\rho_{XY} = 0$
- $Y = aX + b$ , gdje je  $a > 0$ , onda i samo onda ako je  $\rho_{XY} = 1$
- $Y = aX + b$ , gdje je  $a < 0$ , onda i samo onda ako je  $\rho_{XY} = -1$ .

Ako je, kažemo da su slučajne varijable  $X$  i  $Y$  nekolinearne.

Navedena svojstva koeficijenta korelacije upućuju na činjenicu da zavisnost između slučajnih varijabli  $X$  i  $Y$  možemo potvrditi ako pokažemo da je njihov koeficijent korelacije različit od 0. Osim toga, ako je koeficijent korelacije 1 ili -1, onda znamo i tip veze između  $X$  i  $Y$ , tj. u tim slučajevima ta je veza linearna (Benšić, Šuvak, 2013).

## 2.5. KEMOMETRIJSKE METODE

Kemometrija se definira kao kemijska disciplina koja koristi matematičke i statističke metode, za oblikovanje ili odabir optimalnog mjeriteljskog postupaka ili eksperimenta i za mogućnost dobivanja maksimalnog broja informacija analizom dobivenih podataka (Mathias 2007).

Kemometrijske tehnike koje su korištene pri obradi podataka u ovom radu su analiza glavnih komponenta (eng. *principal component analysis, PCA*), klaster analiza (eng. *cluster analysis, CA*) i regresijska analiza metodom najmanjih kvadrata (eng. *partial least square regression, PLS*).

### 2.5.1. Analiza glavnih komponenti (PCA)

Analiza glavnih komponenta je statistička metoda koja se koristi kod obrade podataka s više varijabli. Cilj metode je objasniti što veći broj varijabli pomoću što manjeg broja varijabli formiranjem novih, umjetnih varijabli koje su linearne složenice – kombinacije izvornih varijabli. Maksimalni broj novih varijabli koji se može formirati jednak je broju izvornih, a nove varijable nisu međusobno korelirane (Pecina, 2006).

Ovom metodom se veći broj visoko koreliranih varijabli zamjenjuje manjim brojem varijabli.

Metoda stvara  $p$  linearnih kombinacija izvornih varijabli koje se nazivaju glavne komponente (eng. *principal components*):

$$\begin{aligned}\xi_1 &= w_{11} X_1 + w_{12} X_2 + \dots + w_{1p} X_p \\ \xi_2 &= w_{21} X_1 + w_{22} X_2 + \dots + w_{2p} X_p \\ &\vdots \quad \vdots \quad \quad \quad \vdots \quad \vdots \quad \vdots \\ \xi_p &= w_{p1} X_1 + w_{p2} X_2 + \dots + w_{pp} X_p\end{aligned}$$

(6)

gdje su  $\xi_1, \xi_2 \dots \xi_p$ ,  $p$  glavnih komponenta i  $w_{ij}$  su težinski koeficijenti (eng. *weights*) tj. konstante koje čine koeficijente  $j$ -te varijable za  $i$ -tu glavnu komponentu.

Konstante  $w_{ip}$  nazivaju se svojstveni ili latentni vektori (eng. *eigenvectors*) i geometrijski su u dvodimenzionalnoj strukturi.



Suma varijanci svih izvornih varijabli je ukupna varijanca, dio te varijance objašnjen jednom glavnom komponentom naziva se svojstvena vrijednost ili latentni korijen (eng. *eigenvalue*).

Svojstvena vrijednost najveća je u prvoj komponenti i u svakoj sljedećoj njena vrijednost je manja. Suma svih svojstvenih vrijednosti jednaka je ukupnoj varijanci.

Cilj ove metode je izdvojiti čim veći dio ukupne varijance u tek nekoliko prvih glavnih komponenti, što se izražava u kumulativnim postocima ukupne varijance i time se smanjuje broj izvornih varijabli.

### **2.5.2. Klaster analiza (CA)**

Klaster analiza je grupa multivarijantnih čiji je primarni cilj klasificiranje ili klasteriranje opažaja u skupine, grupe ili klastere (Pecina, 2006).

Klaster analiza nije metoda statističkog zaključivanja uz pretpostavku odnosa uzorak-populacija, nego objektivna metodologija za klasificiranje. Ova analiza je osjetljiva na izbor varijabli relevantnih za objekt istraživanja.

### **2.5.3. Regresijska analiza metodom najmanjih kvadrata (PLS)**

Regresijska analiza metodom najmanjih kvadrata je metoda zasnovana na višestrukoj regresiji i analizi glavnih komponenata. Može se koristiti na multikolinearnim podacima, može uključivati veliki skup nezavisnih varijabli, a osim toga nekoliko zavisnih varijabli može se modelirati istovremeno (Piggot, 1986). PLSR metoda određuje linearnu kombinaciju zavisnih varijabli. Najdjelotvornija je kod predviđanja. Cilj korištenja ove metode je predviđanje budućih vrijednosti zavisne varijable kao reakcije na promjene vrijednosti nekoliko nezavisnih varijabli.

### **3. EKSPERIMENTALNI DIO**

### 3.1. ZADATAK

Zadatak ovog rada je odrediti povezanost svojstava pšeničnog zrna i brašna. Laboratorijske analize provedene su na 24 uzorka brašna proizvedena od 24 kultivara pšenice (S. Prolific, U1, Dubrava, Libellula, Bezostaja 1, Zlatna Dolina, Tena, OS Crvenka, Osječka 20, Osječanka, Žitarka, Srpanjka, Demetra, Golubica, S. Žitarka, Lucija, Alka, Sana, Ficko, Renata, Katarina, Felix, Divana). U radu su analizirana sljedeća svojstva: prinos, hektolitarska masa, masa tisuću zrna, udio proteina, sedimentacija, vlažni gluten, gluten indeks, broj padanja i izbrašnjavanje.

### 3.2. MATERIJAL I METODE

Kultivari pšenice korišteni u ovom radu uzgojeni su na poljoprivrednim površinama Poljoprivrednog instituta Osijek u vremenskom rasponu od 10 godina (2005.-2014.). Podaci za svojstva pšeničnog zrna dobivena su od proizvođača, a ostala su svojstva analizirana u laboratoriju Poljoprivrednog Instituta Osijek.

Pšenična zrna od 24 kultivara samljevena su, a 24 uzorka brašna podijeljena su u vrećice i korištena za analize.

#### 3.2.1. Određivanje hektolitarske mase

Hektolitarska masa (kg) određena je prema *HRN ENG ISO 7971-2:2010*.

#### 3.2.2. Određivanje apsolutne mase pšenice

Apsolutna masa pšenice izražava se kao masa tisuću prosječnih zrna izraženih u gramima. Od ispitnog uzorka, uređajem za brojenje ili ručno, odbroji se dva puta po 500 zrna bez primjesa, izvaže s točnošću 0,1 g i vrijednosti se zbroje. Metoda je odrađena prema *HRN ENG ISO 520:2012*.

#### 3.2.3. Određivanje udjela proteina

Količina proteina u cijelom zrnu određena je *Infratec 1241 Grain Analyzer-om (Foss Tecator AB, Švedska)*. *Infratec 1241 Grain Analyzer* je instrument koji radi na principu bliske infracrvene transmisije i ima mogućnost istovremenog, preciznog određivanja nekoliko sastojaka u cijelom zrnu.

### 3.2.4. Određivanje sedimentacijske vrijednosti po Zeleny-u

Metoda se temelji na sposobnosti glutena da bubri pod utjecajem mliječne kiseline. Princip metode je u usitnjavanju i prosijavanju uzoraka pšenice, a potom suspendiranje u otopini mliječne kiseline određeno vrijeme. Suspenzija se taloži i nakon određenog vremena očita se volumen taloga koji predstavlja sedimentacijsku vrijednost, koja se izražava u  $\text{cm}^3$ . Metoda je odrađena prema *HRN EN ISO 5529:2010*.

### 3.2.5. Određivanje vlažnog glutena

Količina i kakvoća glutena određena je metodom *ICC No. 155* na uređaju *Glutomatic 2200 System i Glutomatic Centrifuge 2015*, Perten, Švedska.

Količina vlažnog glutena određuje se ispiranjem (5 minuta) pšeničnog brašna (10 g) 2%-tnom otopinom NaCl. Nakon ispiranja gluten se centrifugira na specijalnom situ (88 mikrona) koje zadržava i dio glutena koji zaostaje na situ i dio koji prolazi.

O kvaliteti vlažnog glutena ovise najvažnije osobine tijesta kao što su: rastezljivost, elastičnost, sposobnost, zadržavanje plina i dr. (Kemija i tehnologija žitarica, 2010/2011).

Udio vlažnog glutena izračuna se prema izrazu:

$$\text{udio vlažnog glutena} = \frac{UG * 100}{m} [\%] \quad (7)$$

UG – ukupni gluten (g)

m – masa brašna (g)

Gluten indeks se može izračunati prema izrazu:

$$GI = \frac{GZS (g)}{UG (g)} * 100 \quad (8)$$

gdje je:

GI – gluten indeks

GZS – gluten zaostao na situ (g)

UG – ukupni gluten (g)

Vrijednost gluten indeksa kreće se od 0 do 100. Kultivari sa slabim (rastezljivim ili mekanim) glutenom imaju nizak gluten indeks, za razliku od kultivara s visokim gluten indeksom koje karakterizira jak (kratak) gluten. Optimalna vrijednost gluten indeksa krušnih pšenica kreće se od 60 do 90.

### **3.2.6. Određivanje amilolitičke aktivnosti brašna metodom broja padanja po Hagberg-Pertenu**

Određivanje broja padanja se provodi na uređaju po Hagberg-Pertenu, proizvođač Perten Instruments, Švedska, određena je metodom *ICC No. 107/1*.

Broj padanja (eng. *falling number*) je međunarodna standardna metoda za određivanje aktivnosti  $\alpha$ -amilaze u pšenici i brašnu. Definira se kao ukupno vrijeme od trenutka ulaganja kivete sa suspenzijom u vodenu kupelj, pa do kraja penetracije mješalice viskozimetra kroz škrobni gel. Broj padanja izražava se u sekundama.

Ako je broj padanja manji od 150 sekundi, pšenica je proklijala i aktivnost  $\alpha$ -amilaze je visoka, ako je broj između 200-300 sekundi aktivnost  $\alpha$ -amilaze je normalna. Optimalan raspon vrijednosti broja padanja je između 250 i 300 sekundi.

### **3.2.7. Statistička obrada podataka**

Na mjerenom setu podataka provedena je osnovna statistička analiza određivanja srednjih vrijednosti i medijana, standardne devijacije, koeficijenta varijabilnosti, te koeficijenta korelacije, zatim su primijenjene različite kemometrijske metode: analiza glavnih komponenata (PCA), klasterka analiza (CA) te regresijska metoda najmanjih kvadrata (PLSR). Za statističku obradu podataka korišten je programski paket *Statistica ver. 12*. Deskriptivna statistička analiza korištena je za izračunavanje promjena veličina tijekom analiziranog vremenskog razdoblja i opisivanje prirode tih promjena. Kemometrijske analize korištene su za određivanje najmanjeg broja varijabli koje mogu biti upotrijebljene za opisivanje varijance, za grupiranje svojstava koja su visokokorelirana i razdvajanje svojstava s niskom

korelacijom (PCA), za grafičko prikazivanje međusobne povezanosti svojstava (CA) i izradu prediktivnih matematičkih modela za izračunavanje nekih svojstava pšenice i brašna (PLSR).

## **4. REZULTATI**

#### 4.1. DESKRIPTIVNA STATISTIKA

Na izmjerenom setu podataka u razdoblju od 10 godina (2005.-2014., **Tablica 1**) provedena je deskriptivna statistička analiza i određivane su srednje vrijednosti, medijani, standardne devijacije (Std. dev.) i koeficijenti varijabilnosti za desetogodišnje vrijednosti mjerenih veličina (Koef. var., %), (**Tablica 2**).

**Tablica 1** Izmjerene srednje vrijednosti svojstava pšeničnog zrna i brašna od 24 kultivara

H	Y	HL	TKW	P	SED	WG	GI	FN	FY
2005	6675,526	77,382	38,597	15,383	58,292	37,533	78,081	237,958	75,433
2006	8269,987	83,679	43,321	15,021	54,542	36,863	75,438	344,750	71,250
2007	6538,872	81,432	37,553	15,117	53,125	35,671	76,437	247,917	68,042
2008	8817,821	83,253	43,272	13,092	38,417	29,029	88,575	291,125	67,917
2009	8497,921	80,024	45,390	13,733	46,625	29,108	88,820	305,625	67,458
2010	6352,669	76,893	33,917	15,888	49,458	37,440	82,948	305,125	72,208
2011	8538,092	84,765	43,663	12,550	38,375	29,094	88,289	344,708	68,833
2012	8065,523	82,740	40,720	11,546	26,458	25,433	85,845	323,000	69,125
2013	6225,017	76,201	33,742	14,500	49,792	33,708	78,264	410,750	61,542
2014	4190,989	72,441	31,991	14,075	41,667	35,221	77,710	366,688	63,167

H = eng. *harvest*, Y = eng. *yield (kg ha<sup>-1</sup>)*, HL = eng. *hectolitre (kg)*, TKW = eng. *thousand kernel weight (g)*, P = eng. *protein (%)*, SED = eng. *sedimentation (cm<sup>3</sup>)*, WG = eng. *wet gluten (%)*, GI = eng. *gluten index*, FN = eng. *falling number (s)*, FY = eng. *flour yield (%)*

**Tablica 2** Deskriptivna statistička analiza svojstava pšeničnog zrna i brašna (N=10)

Varijabla	Medijan	Mean	Minimum	Maksimum	Std.dev.	Koef.var.
Y	7370,525	7217,242	4190,989	8817,821	1468,874	20,35230
HL	80,728	79,881	72,441	84,765	4,004	5,01286
TKW	36,659	39,217	31,991	45,390	4,794	12,22467
P	14,288	14,090	11,546	15,888	1,373	9,74267
SED	48,042	45,675	26,458	58,292	9,526	20,85709
WG	34,465	32,910	25,433	37,533	4,360	13,24858
GI	80,606	82,041	75,438	88,820	5,444	6,63518
FN	314,313	317,765	237,958	410,750	52,478	16,51488
FY	68,438	68,498	61,542	75,433	4,060	5,92788

H = eng. *harvest*, Y = eng. *yield (kg ha<sup>-1</sup>)*, HL = eng. *hectolitre (kg)*, TKW = eng. *thousand kernel weight (g)*, P = eng. *protein (%)*, SED = eng. *sedimentation (cm<sup>3</sup>)*, WG = eng. *wet gluten (%)*, GI = eng. *gluten index*, FN = eng. *falling number (s)*, FY = eng. *flour yield (%)*

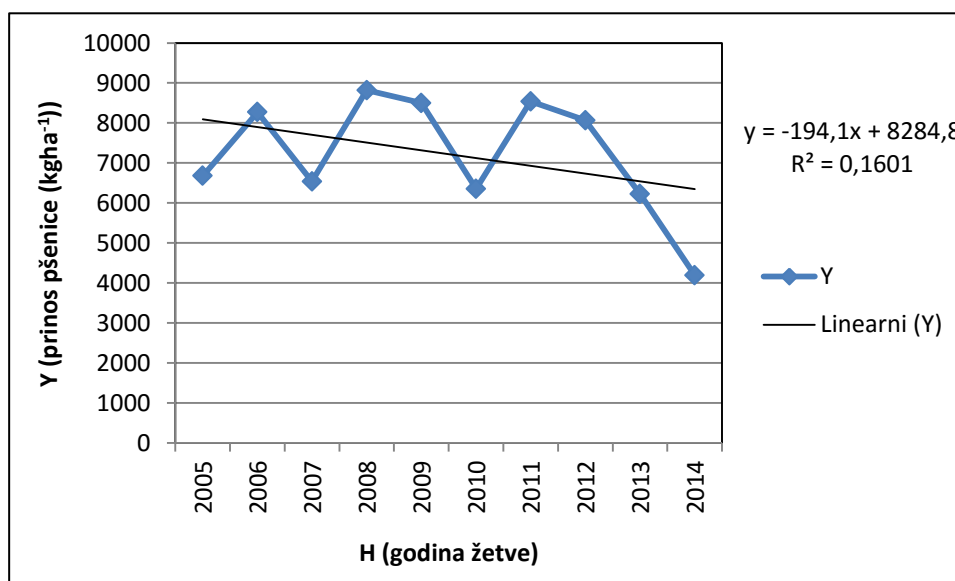


Linearne regresijske jednadžbe i koeficijenti determinacije za promjene analiziranih svojstava dani su u **Tablici 3**, te primjer grafičkog prikaza promjene i linearne zavisnosti (trenda promjene) mjerene veličine tijekom analiziranog razdoblja (**Slika 2**).

**Tablica 3** Jednadžbe linearne regresije i koeficijenti determinacije za promjene analiziranih svojstava

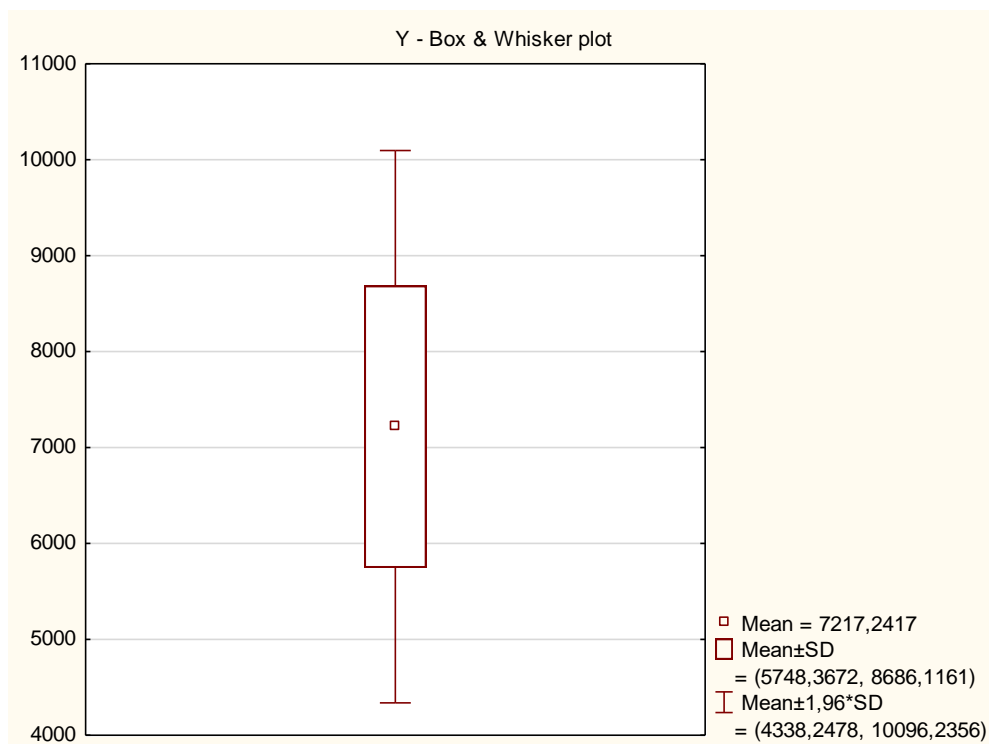
Svojstvo	Jednadžba linearne regresije	R <sup>2</sup>
Y	$y = -194,1x + 8284,8$	0,1601
HL	$y = -0,5386x + 82,843$	0,1658
TKW	$y = -0,7331x + 43,249$	0,2144
P	$y = -0,1985x + 15,182$	0,1916
SED	$y = -1,9x + 56,125$	0,3646
WG	$y = -0,5185x + 35,762$	0,1296
GI	$y = 0,3439x + 80,149$	0,0366
FN	$y = 13,068x + 245,89$	0,5684
FY	$y = -1,0027x + 74,012$	0,559

H = eng. *harvest*, Y = eng. *yield* ( $\text{kg ha}^{-1}$ ), HL = eng. *hectolitre* (kg), TKW = eng. *thousand kernel weight* (g), P = eng. *protein* (%), SED = eng. *sedimentation* ( $\text{cm}^3$ ), WG = eng. *wet gluten* (%), GI = eng. *gluten index*, FN = eng. *falling number* (s), FY = eng. *flour yield* (%)

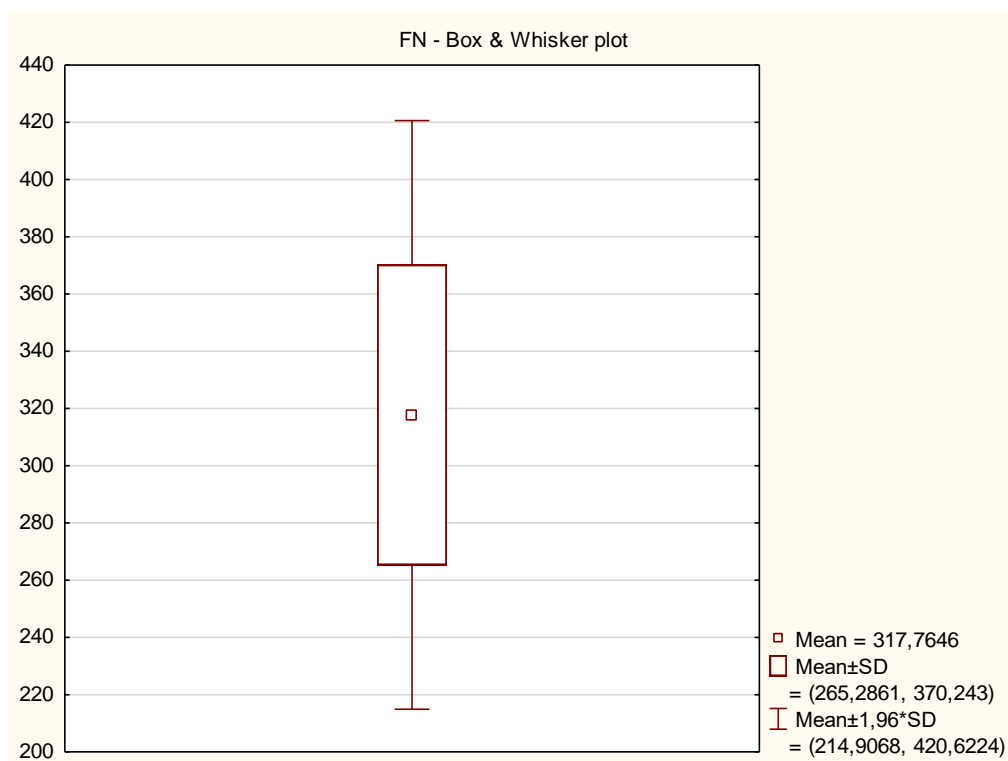


**Slika 2** Mjerene vrijednosti i trend promjene prinosa pšenice za razdoblje 2005.-2014. godine

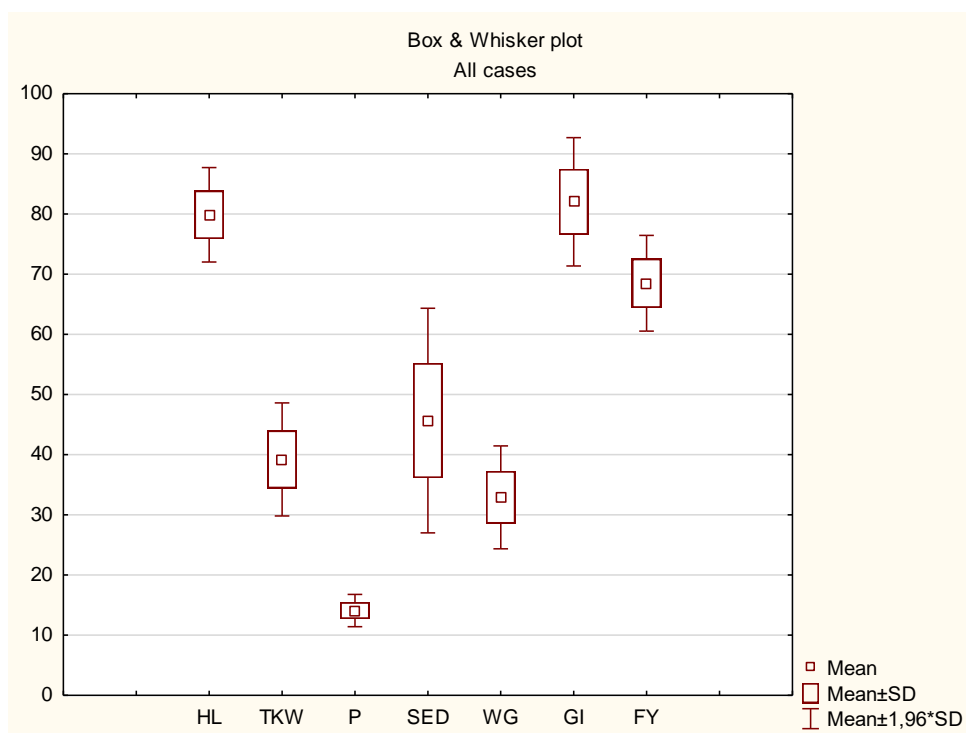
**Slike 3-5** prikazuju srednje vrijednosti, raspone promjena i standardne devijacije mjerenih veličina za razdoblje 2005.-2014. godine, dobivene korištenjem programskog paketa *Statistica ver. 12*.



**Slika 3** Box & Whisker dijagram za prinos pšenice Y (kg ha<sup>-1</sup>)



**Slika 4** Box & Whisker dijagram za vrijednost broja padanja FN (s)



**Slika 5** Box & Whisker dijagram za ostala svojstva pšenice i brašna

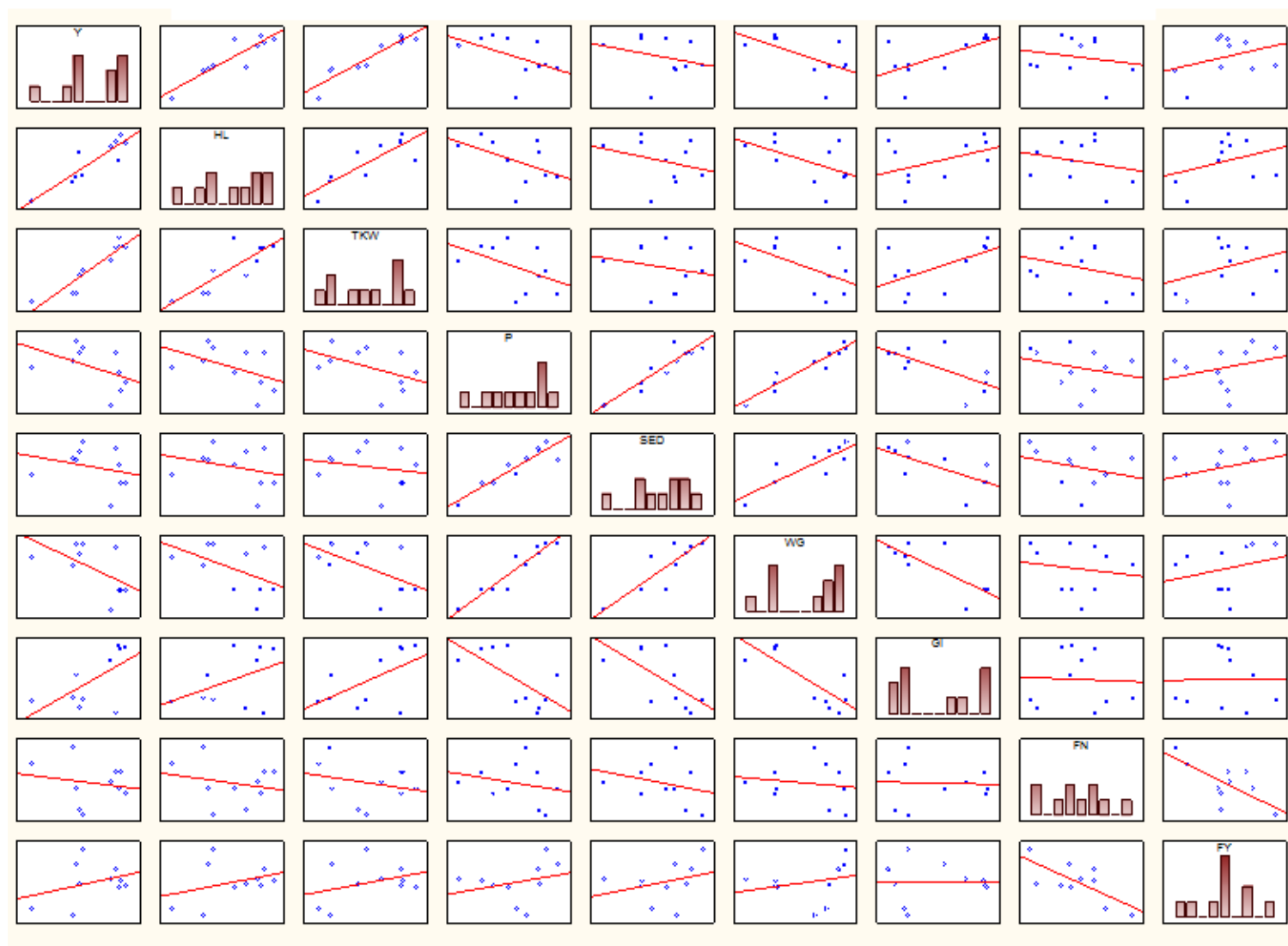
(hektolitarska masa HL (kg), masa tisuću zrna TKW (g), udio bjelančevina P (%), sedimentacijska vrijednost SED (cm<sup>3</sup>), vlažni gluten WG (%), gluten indeks GI, izbrašnjavanje FY (%))

U **Tablici 4** prikazani su koeficijenti korelacije svojstava kvalitete pšeničnog zrna i brašna, a na **Slici 6** dan je prikaz linearne zavisnosti svojstava temeljen na koeficijentu korelacije.

**Tablica 4** Koeficijenti korelacije za svojstva kvalitete pšeničnog zrna i brašna od 24 kultivara

Varijabla	Y	HL	TKW	P	SED	WG	GI	FN	FY
Y	1,00	0,90	0,93	-0,47	-0,27	-0,60	0,63	-0,19	0,34
HL	0,90	1,00	0,84	-0,47	-0,29	-0,51	0,43	-0,22	0,33
TKW	0,93	0,84	1,00	-0,46	-0,20	-0,55	0,57	-0,25	0,33
P	-0,47	-0,47	-0,46	1,00	0,92	0,95	-0,67	-0,25	0,28
SED	-0,27	-0,29	-0,20	0,92	1,00	0,86	-0,65	-0,28	0,30
WG	-0,60	-0,51	-0,55	0,95	0,86	1,00	-0,81	-0,15	0,26
GI	0,63	0,43	0,57	-0,67	-0,65	-0,81	1,00	-0,04	0,02
FN	-0,19	-0,22	-0,25	-0,25	-0,28	-0,15	-0,04	1,00	-0,71
FY	0,34	0,33	0,33	0,28	0,30	0,26	0,02	-0,71	1,00

H = eng. harvest, Y = eng. yield (kg ha<sup>-1</sup>), HL = eng. hectolitre (kg), TKW = eng. thousand kernel weight (kg), P = eng. protein (%), SED = eng. sedimentation (cm<sup>3</sup>), WG = eng. wet gluten (%), GI = eng. gluten index (%), FN = eng. falling number (s), FY = eng. flour yield (%)



**Slika 6** Prikaz linearnih zavisnosti na temelju koeficijenata korelacije

## 4.2. KEMOMETRIJSKE METODE

### 4.2.1. Analiza glavnih komponenti (PCA)

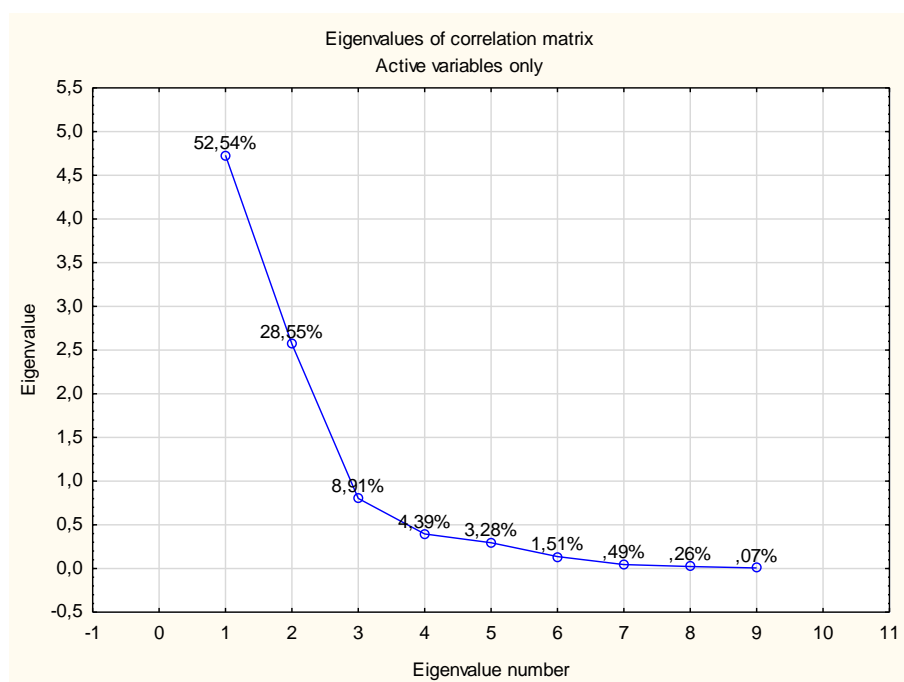
Na izmjerenom setu podataka za razdoblje od 10 godina (2005-2014, **Tablica 1**) provedena je i kemometrijska statistička analiza.

U **Tablici 5** prikazane su svojstvene vrijednosti korelacijske matrice. Na **Slici 7** dan je grafički prikaz iznosa ukupne varijance, **Slike 8 i 9** prikazuju raspodjelu u prve dvije faktorske ravnine, koje zajedno opisuju 81,09% ukupne varijance.

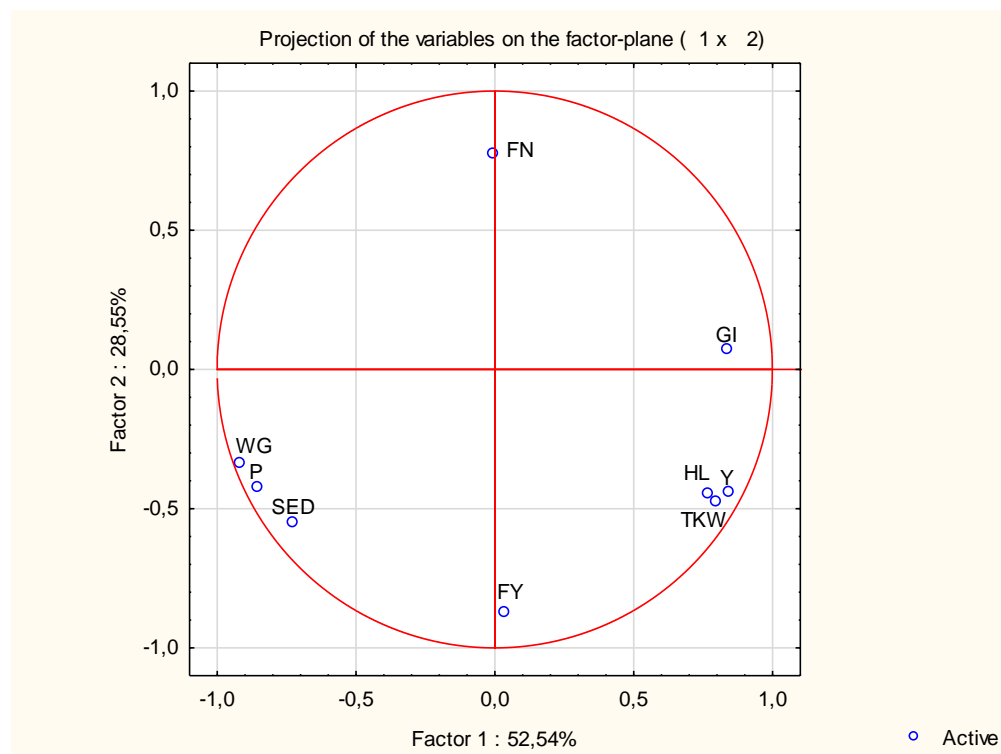
**Tablica 5** Svojstvene vrijednosti korelacijske matrice i pridruženi pokazatelji

Broj vrijednosti	Svojstvena vrijednost	Udio ukupne varijance (%)	Kumulativna svojstvena vrijednost	Kumulativna varijanca (%)
1	4,728887	52,54319	4,728887	52,5432
2	2,569298	28,54776	7,298185	81,0909
3	0,801926	8,91029	8,100111	90,0012
4	0,395313	4,39237	8,495424	94,3936
5	0,294820	3,27578	8,790245	97,6694
6	0,136093	1,51214	8,926337	99,1815
7	0,044458	0,49398	8,970795	99,6755
8	0,023325	0,25916	8,994120	99,9347
9	0,005880	0,06534	9,000000	100,0000

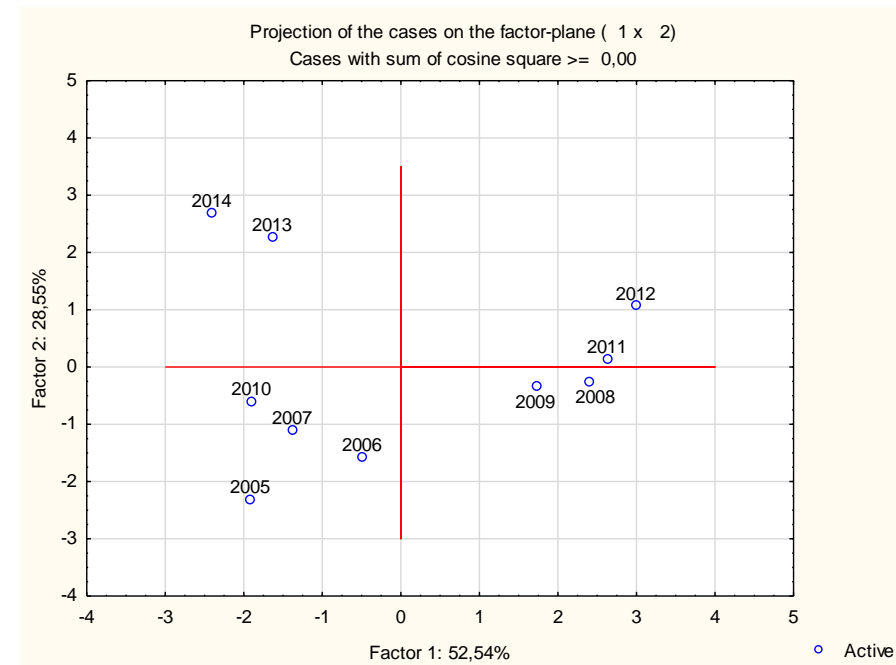
H = eng. *harvest*, Y = eng. *yield* ( $\text{kg ha}^{-1}$ ), HL = eng. *hectolitre* (kg), TKW = eng. *thousand kernel weight* (g), P = eng. *protein* (%), SED = eng. *sedimentation* ( $\text{cm}^3$ ), WG = eng. *wet gluten* (%), GI = eng. *gluten index*, FN = eng. *falling number* (s), FY = eng. *flour yield* (%)



**Slika 7** Iznosi udjela ukupne varijance glavnih komponenata



**Slika 8** Grupiranje svojstava u faktorskoj ravlini prve dvije glavne komponente iz analize glavnih komponentata

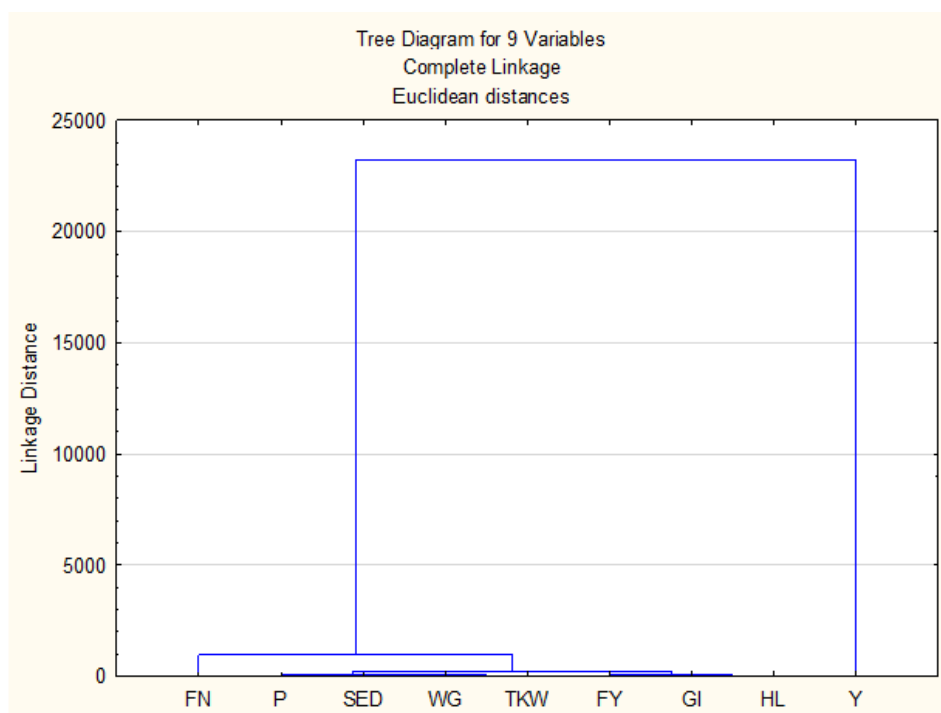


**Slika 9** Grupiranje godina žetve u faktorskoj ravlini prve dvije glavne komponente iz analize glavnih komponentata

### 4.2.2. Klaster analiza (CA)

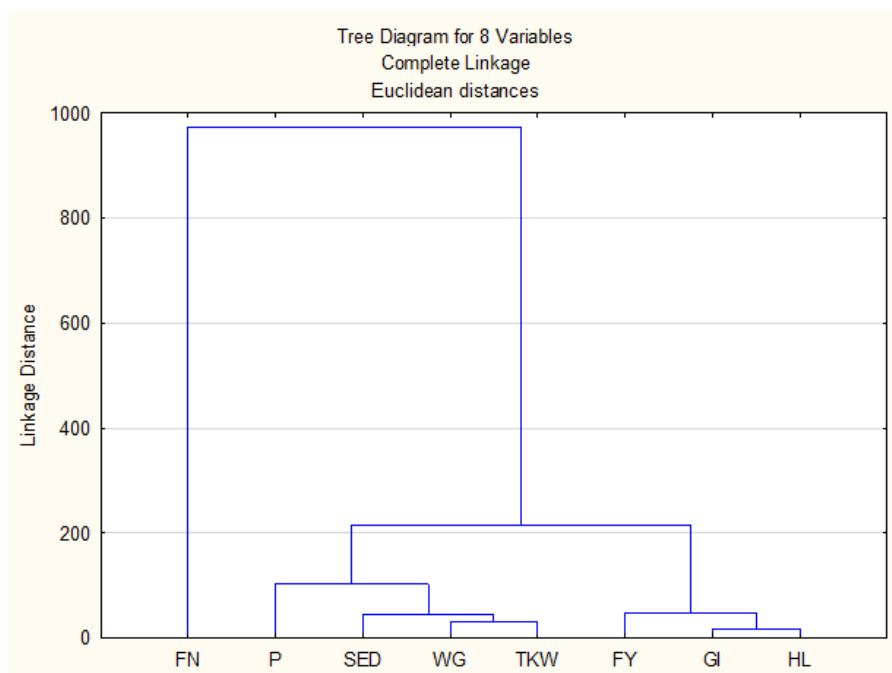
Primarni cilj klaster analize je klasificiranje ili klasteriranje opažaja u skupine, grupe ili klastere tako da je svaka grupa ili klaster kompaktan obzirom na određene varijable (svojstva) ili svaka grupa različita od druge obzirom na te iste varijable (svojstva).

Na **Slikama 10** i **11** dani su dendogrami klaster analize svojstava pšeničnog zrna i brašna.



**Slika 10** Dendogram klaster analize sa svim svojstvima





**Slika 11** Dendrogram klaster analize bez prinosa (Y)

### 4.2.3. Regresijska analiza metodom najmanjih kvadrata (PLSR)

Metoda PLS korištena je u **Modelu 1-5**, gdje se pomoću različitih nezavisnih varijabli izračunava vrijednost jedne zavisne varijable. Validacija predloženih prediktivnih modela provedena je na temelju koeficijenta korelacije između vrijednosti izmjerene varijable i vrijednosti dobivene modelom.

#### 4.2.3.1. Model 1 - PLSR prediktivni model za izračunavanje TKW pomoću Y i HL

U **Modelu 1** nezavisne varijable Y i HL upotrijebljene su za izračunavanje vrijednosti zavisne varijable TKW.

**Tablica 6** PLS regresijski koeficijenti

	Interc.	Y	HL
TKW	17,18401	0,003048	0,000414

Y = eng. *yield* (kg/ha-1) , HL = eng. *hectolitre* (kg), TKW = eng. *thousand kernel weight* (g), Interc. = eng. *interception* (slobodni član u jednadžbi predstavlja sjecište regresijskog pravca i ordinatne osi)

Izmjerene vrijednosti Y i HL za 2005. godinu bile su 6675,526 i 77,382 (**Tablica 1**). Na temelju svih vrijednosti za razdoblje od 2005.-2014. godine određeni su regresijski koeficijenti prediktivnog modela za izračunavanje vrijednosti varijable TKW (**Tablica 6**).

Primjer izračuna:

$$\text{TKW} = 0,003048 * Y + 0,000414 * \text{HL} + 17,18401 \quad (9)$$

Korištenjem **Modela 1** temeljenog na jednadžbi (9) izračunate su vrijednosti mase tisuću zrna i određen je koeficijent korelacije između eksperimentalnih (izmjerenih) i prediktivnim **Modelom 1** izračunatih vrijednosti (**Tablica 7**).

**Tablica 7** Korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti

H	TKW_mj	TKW_model	Koef. korel.
2005	38,597	37,56427	0,934251
2006	43,321	42,42709	
2007	37,553	37,14940	
2008	43,272	44,09681	
2009	45,390	43,12036	
2010	33,917	36,57994	
2011	43,663	43,24477	
2012	40,720	41,80346	
2013	33,742	36,19055	
2014	31,991	29,98890	

#### 4.2.3.2. Model 2 - PLSR prediktivni model za izračunavanje P pomoću Y, HL i TKW

U **Modelu 2** nezavisne varijable Y, HL i TKW upotrijebljene su za izračunavanje vrijednosti zavisne varijable P.

**Tablica 8** PLS regresijski koeficijent

	Interc.	Y	HL	TKW
P	22,64666	-0,000130	-0,074550	-0,042337

P = eng. *protein* (%) Y = eng. *yield* (kg/ha-1), HL = eng. *hectolitre* (kg), TKW = eng. *thousand kernel weight* (g), Interc. = eng. *interception* (slobodni član u jednadžbi predstavlja sjecište regresijskog pravca i ordinatne osi)

Izmjerene vrijednosti Y, HL i TKW za 2005. godinu bile su 6675,526,77,382 i 38,597 (**Tablica 1**). Na temelju svih vrijednosti za razdoblje od 2005.-2014. godine određeni su regresijski koeficijenti prediktivnog modela za izračunavanje vrijednosti varijable TKW (**Tablica 8**).

Primjer izračuna:

$$P = -0,000130 \cdot Y + (-0,074550) \cdot HL + (-0,042337) \cdot TKW + 22,64666 \quad (10)$$

Korištenjem **Modela 2** temeljenog na jednadžbi (10) izračunate su vrijednosti mase tisuću zrna i određen je koeficijent korelacije između eksperimentalnih (izmjerenih) i prediktivnim **Modelom 2** izračunatih vrijednosti (**Tablica 9**).

**Tablica 9** Korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti

H	P_mj	P_model	Koef. korel.
2005	15,383	14,37354	0,485273
2006	15,021	13,49627	
2007	15,117	14,13367	
2008	13,092	13,45869	
2009	13,733	13,65149	
2010	15,888	14,65025	
2011	12,550	13,36588	
2012	11,546	13,70304	
2013	14,500	14,72589	
2014	14,075	15,34543	

#### 4.2.3.3. Model 3 - PLSR prediktivni model za izračunavanje GI pomoću P, SED i WG

U **Modelu 3** nezavisne varijable P, SED i WG upotrijebljene su za izračunavanje vrijednosti zavisne varijable GI.

**Tablica 10** PLS regresijski koeficijent

	Interc.	P	SED	WG
GI	90,92969	4,957721	-0,190449	-2,12843

P = eng. *Protein*, SED = eng. *sedimentation* (cm<sup>3</sup>) , WG = eng. *wet gluten* (%) , GI = eng. *gluten index*, Interc. = eng. *interception* (slobodni član u jednadžbi predstavlja sjecište regresijskog pravca i ordinatne osi)

Izmjerene vrijednosti P, SED i WG za 2005. godinu bile 15,383, 58,292 i 37,533 (**Tablica 1**). Na temelju svih vrijednosti za razdoblje od 2005.-2014. godine određeni su regresijski koeficijenti prediktivnog modela za izračunavanje vrijednosti varijable GI (**Tablica 10**).

Primjer izračuna:

$$GI = 4,9577721 \cdot 15,383 + (-0,190449) \cdot 58,292 + (-2,12843) \cdot 37,533 + 90,92969 \quad (11)$$

Korištenjem **Modela 2** temeljenog na jednadžbi (11) izračunate su vrijednosti mase tisuću zrna i određen je koeficijent korelacije između eksperimentalnih (izmjerenih) i prediktivnim **Modelom 3** izračunatih vrijednosti (**Tablica 11**).

**Tablica 11** Korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti

H	GI_mjereno	GI_model	Koef. korel.
2005	78,081	76,20743	0,867923
2006	75,438	76,55226	
2007	76,437	79,83356	
2008	88,575	86,73166	
2009	88,820	88,18154	
2010	82,948	80,58791	
2011	88,289	83,91581	
2012	85,845	88,99877	
2013	78,264	81,58861	
2014	77,710	77,80840	

#### 4.2.3.4. Model 4 - PLSR prediktivni model za izračunavanje GI pomoću P, SED, WG i FN

U **Modelu 4** nezavisne varijable P, SED, WG i FN upotrijebljene su za izračunavanje vrijednosti zavisne varijable GI.

**Tablica 12** PLS regresijski koeficijent

	Interc.	P	SED	WG	FN
GI	96,11882	4,670634	-0,209650	-2,02764	-0,011278

P = eng. *Protein*, SED = eng. *sedimentation* (cm<sup>3</sup>), WG = eng. *wet gluten* (%), GI = eng. *gluten index* (%), FN = eng. *falling number*, Interc. = eng. *interception* (slobodni član u jednadžbi predstavlja sjecište regresijskog pravca i ordinatne osi)

Izmjerene vrijednosti P, SED, WG i FN za 2005. godinu bile su 15,383, 58,292, 37,533 i 237,958 (**Tablica 1**). Na temelju svih vrijednosti za razdoblje od 2005.-2014. godine određeni su regresijski koeficijenti prediktivnog modela za izračunavanje vrijednosti varijable GI (**Tablica 12**).

Primjer izračuna:

$$GI = 4,670634 * 15,383 + (-0,209650) * 58,292 + (-2,02764) * 37,533 + (-0,011278) * 237,958 + 96,11882 \quad (12)$$

Korištenjem **Modela 4** temeljenog na jednadžbi **11** izračunate su vrijednosti mase tisuću zrna i određen je koeficijent korelacije između eksperimentalnih (izmjerenih) i prediktivnim **Modelom 4** izračunatih vrijednosti (**Tablica 13**).

**Tablica 13** Korelacije između izmjerenih i PLSR modelom izračunatih vrijednost

H	GI_mjereno	GI_model	Koef. korel.
2005	78,081	76,96001	0,873789
2006	75,438	76,20890	
2007	76,437	80,46188	
2008	88,575	87,06706	
2009	88,820	88,01954	
2010	82,948	80,59844	
2011	88,289	83,80976	
2012	85,845	89,28568	
2013	78,264	80,42371	
2014	77,710	77,57096	

#### 4.2.3.5. Model 5 - PLSR prediktivni model za izračunavanje GI pomoću P, SED, WG i FY

U **Modelu 5** nezavisne varijable P, SED, WG i FY upotrijebljene su za izračunavanje vrijednosti zavisne varijable GI.

**Tablica 14** PLS regresijski koeficijent

	Interc.	P	SED	WG	FY
GI	71,73359	4,934387	-0,228494	-2,12362	0,308103

P = eng. *Protein*, SED = eng. *sedimentation* (cm<sup>3</sup>) , WG = eng. *wet gluten* (%) , GI = eng. *gluten index*, FY = eng. *flour yield* (%), Interc. = eng. *interception* (slobodni član u jednadžbi predstavlja sjecište regresijskog pravca i ordinatne osi)

Izmjerene vrijednosti P, SED, WG i FY za 2005. godinu bile 15,383, 58,292, 37,533 i 75,433 (**Tablica 1**). Na temelju svih vrijednosti za razdoblje od 2005.-2014. godine određeni su regresijski koeficijenti prediktivnog modela za izračunavanje vrijednosti varijable GI (**Tablica 14**).

Primjer izračuna:

$$GI = 4,934387 \cdot 15,383 + (-0,228494) \cdot 58,292 + (-2,12362) \cdot 37,533 + 0,308103 \cdot 75,433 + 71,73359 \quad (13)$$

Korištenjem **Modela 5** temeljenog na jednadžbi (13) izračunate su vrijednosti mase tisuću zrna i određen je koeficijent korelacije između eksperimentalnih (izmjerenih) i prediktivnim **Modelom 5** izračunatih vrijednosti (**Tablica 15**)

**Tablica 15** Korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti

H	GI_mjereno	GI_model	Koef. korel.
2005	78,081	77,85646	0,895205
2006	75,438	77,06029	
2007	76,437	79,39901	
2008	88,575	86,83348	
2009	88,820	87,81527	
2010	82,948	81,56714	
2011	88,289	84,31460	
2012	85,845	89,94661	
2013	78,264	79,28316	
2014	77,710	76,32993	

## **5. RASPRAVA**

U radu su analizirana svojstva pšenice i brašna. Na eksperimentalnim vrijednostima za 24 kultivara pšenice i brašna dobivena njihovim mljevenjem, dobivenim tijekom razdoblja 2005.-2014. godine (**Tablica 1**) provedena je deskriptivna statistička analiza (**Tablica 2**), te su provedene različite kemometrijske analize podataka.

Deskriptivnom statističkom analizom određena je srednja vrijednost, medijan, minimalna i maksimalna vrijednost, standardna devijacija i koeficijent varijabilnosti. Najveću promjenu vrijednosti tijekom analiziranog razdoblja imao je prinos pšenice (Y) za koji je koeficijent varijabilnosti iznosio 20,35%, a najmanju varijabilnost od svega 5,01% imala je hektolitarska masa (HL). Najveću varijabilnost (20,86%) među svojstvima brašna imala je sedimentacijska vrijednost (SED), a najmanju varijabilnost (5,93%) imalo je izbrašnjavanje (FY).

Linearnim regresijskim jednadžbama i koeficijentima determinacije (**Tablica 3**) matematički su opisane promjene vrijednosti varijabli tijekom godina i korelacija eksperimentalnih podataka s modelima linearne regresije. Najveću vrijednost koeficijenta determinacije ( $R^2$ ) imali su broj padanja (FN) i izbrašnjavanje (FY). Obje vrijednosti su približno 0,56 te se može zaključiti da su promjene svih analiziranih svojstava nelinearne te slabo koreliraju s regresijskim pravcem. Iz koeficijenata uz varijablu "x" u jednadžbama linearne regresije vidimo da je najveći pad srednjih vrijednosti tijekom razdoblja 2005.-2014. imao prinos, a jedini porast prosječnih vrijednosti je zabilježen kod gluten indeks vrijednosti (GI) i kod broja padanja (FN).

Na **Slici 2** prikazane su mjerene vrijednosti prinosa (Y) tijekom 10 godina i trend njihove promjene.  $R^2$  vrijednost nije značajna, a samim tim ni model nije reprezentativan. Međutim iz podataka možemo zaključiti da je tijekom navedenog vremenskog razdoblja vidljiv trend pada prosječnog prinosa za analizirane kultivare.

Na **Slikama 3-5** (tzv. Box&Whisker dijagrami) prikazane su raspršenosti podataka oko srednjih vrijednosti tijekom razdoblja 2005.-2014. godine, a vidljivo je da je najmanje odstupanje od srednje vrijednosti imala vrijednost udjela proteina (P).

Koeficijenti korelacije za svojstva kvalitete pšeničnog zrna i brašna od 24 kultivara dana su u **Tablici 4**. Svojstva Y, HL i TKW imaju visoku pozitivnu međusobnu korelaciju, što znači da se njihove srednje vrijednosti tijekom razdoblja od deset godina jednako ili slično ponašaju dok



svojstva P, SED, WG imaju negativnu korelaciju s Y, HL i TKW. Iz navedenih korelacija vidljivo je da povećanje prinosa, hektolitarske mase i mase tisuću zrna pojedinog kultivara pšenice ima za posljedicu smanjenje udjela proteina, sedimentacijske vrijednosti i udjela vlažnog glutena u brašnu. Podaci se slažu s literaturnim izvorom (Španić i sur., 2013).

Na **Slici 6** vidljivi su odnosi među analiziranim veličinama i uočljivije su međusobne zavisnosti parova svojstava pšenice i brašna. Korelacijski pravci (crvene linije na dijagramima) pokazuju koje su veličine (svojstva) proporcionalne, a koje obrnuto proporcionalne tijekom razdoblja 2005.-2014. godine.

**Tablica 5** prikazuje svojstvene vrijednosti korelacijske matrice za 9 svojstava 24 kultivara tijekom 10 godina. Svojstvene vrijednosti su varijance glavnih komponenti, njihov zbroj je jednak 9. U drugom stupcu izračunat je postotak ukupne varijance objašnjen svakom glavnom komponentom, a u četvrtom stupcu je dan kumulativni niz postotaka iz drugog stupca. Prvom glavnom komponentom objašnjeno je 52,54% ukupne varijance, drugom 28,54% i trećom komponentom 8,91%, tj. s prve tri glavne komponente objašnjeno je 90,00% ukupne varijance. Grafički prikaz postotka ukupne varijance (tzv. *Scree plot*) dan je na **Slici 7**. Prve dvije glavne komponente opisuju 81,1% ukupne varijance.

Grupiranje svojstava pšeničnog zrna i brašna prikazano je na **Slici 8** u faktorskoj ravnini prve dvije glavne komponente. Grupiranje rezultata (svojstava) može se tumačiti kao pokazatelj sličnosti između međusobno povezanih svojstava. Vidljivo je zajedničko grupiranje Y, HL i TKW koji imaju međusobnu pozitivnu korelaciju, te je nasuprot grupiranje P, SED i WG koji imaju međusobnu pozitivnu korelaciju, a negativnu korelaciju sa Y, HL i TKW. Svojstva FN i FY nalaze se nasuprot jedan drugog, što ukazuje na njihov slab međusobni utjecaj.

Prosječne vrijednosti svih analiziranih svojstava označene godinama žetve grupirane su u faktorskoj ravnini prve dvije komponente prema sličnom ponašanju (**Slika 9**).

Prema **Tablici 1** i **Slici 9** vidljiva je sličnost (mala udaljenost) svojstava 2013. i 2014. godine, koja imaju nizak prinos po hektaru (Y) kojeg prati i nizak udio proteina (P). Grupiranje godina 2005., 2006., 2007. i 2010. dogodilo se zbog visokog udjela proteina (P), vlažnog glutena (WG), veće vrijednost sedimentacije (SED) i nižeg prinosa po hektaru (Y). Godine 2008., 2009., 2011. i 2012. pokazuju visok prinos pšenice (Y) i manji udio proteina (P).

Klusterskom analizom dobiveni su dendogrami (**Slika 10-11**). Iz dendograma se vidi izrazita povezanost svih svojstava s prinosom (Y). Kada se izdvoji vrijednost prinosa vidljiva je povezanost svih mjerenih vrijednosti s vrijednošću broja padanja (FN, **Slika 11**) Svojstva udjela proteina (P), vlažnog glutena (WG), sedimentacijska vrijednost (SED) i masa tisuću zrna (TKW) povezani su grupe, međutim masa tisuću zrna nije u pozitivnoj korelaciji sa proteinom (P), vlažnim glutenom (WG) i sedimentacijskom vrijednošću (SED). Na temelju toga može se zaključiti da su grupe formirane nizom vrijednosti, a ne prema svojstvima.

Regresijskom metodom najmanjih kvadrata (PLSR) izrađeni su prediktivni **Modeli 1-5**. Modeli su matematički povezali analizirana svojstava, a jednadžbe modela prikazane su izrazima (**9-13**).

**Modelom 1** predviđena je vrijednost zavisne varijable mase tisuću zrna pšenice (TKW) iz vrijednosti svojstava nezavisnih varijabli Y i HL (**Tablica 6 i Jednadžba 9**). Uvrštavanjem vrijednosti svojstava Y i HL u jednadžbu **Modela 1** izračunate su vrijednosti varijable TKW. Iz koeficijenta korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti vidljivo je da se s visokom točnošću od 93,4% može predvidjeti vrijednost TKW (**Tablica 7**).

**Modelom 2** predviđena je vrijednost zavisne varijable, proteina (P) iz vrijednosti svojstava nezavisnih varijabli Y, HL i TKW (**Tablica 8 i Jednadžba 10**). Uvrštavanjem vrijednosti svojstava Y, HL i TKW u jednadžbu **Modela 2** izračunate su vrijednosti varijable P. Iz koeficijenta korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti vidljivo je da ovaj model ne odgovara za predviđanje vrijednosti varijable P (**Tablica 9**).

**Modelom 3** predviđena je vrijednost zavisne varijable, gluten indeks (GI) iz vrijednosti svojstava nezavisnih P, WG i SED varijabli (**Tablica 11 i Jednadžba 11**). Uvrštavanjem vrijednosti svojstava P, WG i SED u jednadžbu **Modela 3** izračunate su vrijednosti varijable GI. Iz koeficijenta korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti vidljivo je da se točnošću od 86,8% može predvidjeti vrijednost GI (**Tablica 12**).

**Modelom 4** predviđena je vrijednost zavisne varijable, gluten indeks (GI) iz vrijednosti svojstava nezavisnih varijabli P, WG, SED i FN (**Tablica 12 i Jednadžba 12**). Uvrštavanjem vrijednosti svojstava P, WG, SED i FN u jednadžbu **Modela 4** izračunate su vrijednosti

varijable GI. Iz koeficijenta korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti vidljivo je da se točnošću od 87,4% može predvidjeti vrijednost GI (**Tablica 13**).

**Modelom 5** predviđena je vrijednost zavisne varijable, gluten indeks (GI) iz vrijednosti svojstava nezavisnih varijabli P, WG, SED i FY (**Tablica 13** i **Jednadžba 13**). Uvrštavanjem vrijednosti svojstava P, WG, SED i FY u jednadžbu **Modela 5** izračunate su vrijednosti varijable GI. Iz koeficijenta korelacije između izmjerenih i PLSR modelom izračunatih vrijednosti vidljivo je da se točnošću od 89,5% može predvidjeti vrijednost GI (**Tablica 13**).

Prema dobivenima modelima za predviđanje vrijednosti zavisne varijable GI najboljim se pokazao **Model 5** gdje su nezavisne varijable P, SED, WG i FY. Slična primjena kemometrijskih metoda dana je i u literaturi (Kurtanek i sur., 2003).

## **6. ZAKLJUČCI**

Na temelju rezultata istraživanja provedenih u ovom radu, mogu se izvesti sljedeći zaključci:

- Najveću promjenu vrijednosti tijekom analiziranog razdoblja imao je prinos pšenice (Y) za koji je koeficijent varijabilnosti iznosio 20,35%, a najmanju varijabilnost od svega 5,01% imala je hektolitarska masa (HL). Najveću varijabilnost (20,86%) među svojstvima brašna imala je sedimentacijska vrijednost (SED), a najmanju varijabilnost (5,93%) imalo je izbrašnjavanje (FY).
- Linearnim regresijskim jednadžbama i koeficijentima determinacije matematički su opisane promjene vrijednosti varijabli tijekom godina i korelacija eksperimentalnih podataka s modelima linearne regresije. Promjene svih analiziranih svojstava bile su nelinearne i slabo koreliraju s regresijskim pravcem.
- Koeficijenti korelacije pokazuju da povećanje prinosa (Y), hektolitarske mase (HL) i mase tisuću zrna pojedinog kultivara pšenice (TKW) ima za posljedicu smanjenje udjela proteina (P), sedimentacijske vrijednosti (SED) i udjela vlažnog glutena u brašnu (WG).
- Selekcija kultivara pšenice s visokim prinosima dovodi do smanjenog udjela proteina u pšeničnom brašnu. Na temelju ovoga zaključka za preporučiti je sijati namjensku pšenicu radi proizvodnje optimalne sirovine za prehrambenu industriju i smanjenja potrebe za poboljšivačima u pšeničnom brašnu. Ovakav planski pristup proizvodnji pšenice za prehrambenu industriju značajno smanjuje potrebe za dodatnim radom i troškovima u proizvodnji.
- Primjenom PCA metode je smanjen broj varijabli kojima je moguće opisati većinu promjena analiziranih svojstava. Tako su za predviđanje svojstava pšenice i brašna odabrani prinos pšenice (Y), hektolitarska masa (HL) i masa tisuću zrna pšenice (TKW).
- Kemometrijskim metodama opisano je 90,00% varijance pomoću prve tri glavne komponente, a između tih novih prostornih varijabli vidljiva je visoka korelacija.
- Primjenom klusterske analize dobiveni su dendogrami koji pokazuju grupiranje svojstava koja su međusobno povezani, ali pokazuje i grupiranje svojstava koja nisu međusobno povezana. Na temelju toga može se zaključiti da su grupe formirane nizom vrijednosti, a ne prema svojstvima.

- PLSR metodom kreirani su prediktivni matematički modeli s visokom korelacijom između eksperimentalnih i modelom izračunatih vrijednosti.
- Modelom 1 predviđena je vrijednost masa tisuću zrna pšenice (TKW) s visokom točnošću od 93,4%.
- Model 2 nije odgovarajući za predviđanje vrijednosti udjela proteina (P) zbog niske korelacije između eksperimentalnih i modelom izračunatih vrijednosti.
- Modeli 3, 4 i 5 su predviđali vrijednost gluten indeksa (GI). Najboljim se pokazao model 5 kojim se može predvidjeti vrijednost GI s visokom točnošću od 89,5%.

## **7. LITERATURA**

- Ačkar Đ: *Izoliranje, modificiranje i karakteriziranje škroba pšenice*, Doktorski rad. Prehrambeno-tehnološki fakultet, Osijek, 2010.
- Benšić M, Šuvak N: *Primijenjena statistika*. Sveučilište J.J. Strossmayera, Odjel za matematiku, Osijek, 2013.
- Đaković Lj: *Pšenično brašno*. Tehnološki fakultet, Novi Sad, 1980.
- Hrvatski zavod za norme: *Poljoprivredni i prehrambeni proizvodi; Žito i proizvodi od žita, HRN ENG ISO: 7971-2:2010, ISO 520:2012, 5529:2010*
- International Association for Cereals Science and Technology, ICC Methods, No. 107/1, 155/1
- Kljusurić S: *Uvod u tehnologiju mljevenja pšenice*. Metković, 2000.
- Kurtanek Ž, Horvat D, Jurković Z, Drezner G, Sudar R, Magdić D: *Validation of chemometric analysis of Osijek wheat cultivars*, Food/Kniewald, Medicinska naklada, Zagreb, 2003. str. 185-193
- Mathias O: *Chemometrics*. Wiley – VCH Verlag GmbH & Co. KGaA, Weinheim, 2007.
- Pecina M: *Metode multivarijantne analize-osnove, Interna skripta*. Agronomski fakultet, Sveučilište u Zagrebu, Zagreb, 2006.
- Piggot J.R: *Statistical procedures in food research*. Elsevier Applied Science, London, 1986.
- Pomeranz Y: *Wheat: Chemistry and Technology, Volume II*. American Association of Cereal Chemists, Inc., St. Paul, Minnesota, USA, 1988.
- Prehrambeno-biotehnološki fakultet: *Kemija i tehnologija žitarica, Upute za laboratorijske vježbe*. PBF, Zagreb, 2010./2011.
- Programski paket, *Statistica ver. 12.*, StatSoft Inc, SAD, 2013.
- Španić V, Drezner G, Horvat D, Dvojković K: *Stabilnost parametara kvalitete ozime pšenice u različitim okolinskim uvjetima*, Agronomski glasnik, 2013., str. 2-3
- Ugarčić-Hardi Ž: *Tehnologija proizvodnje i prerade brašna, Opći dio i skladištenje žitarica, Interna skripta*. Prehrambeno-tehnološki fakultet, Osijek, 1999.
- Ugarčić-Hardi Ž: *Tehnologija proizvodnje i prerade brašna, Pekarstvo, Interna skripta*. Prehrambeno-tehnološki fakultet, Osijek, 1999.